

# Package ‘SPLINTER’

October 18, 2017

**Type** Package

**Title** Splice Interpreter Of Transcripts

**Version** 1.2.0

**Date** 2017-03-30

**Author** Diana Low

**Maintainer** Diana Low <dlow@imcb.a-star.edu.sg>

**Description** SPLINTER provides tools to analyze alternative splicing sites, interpret outcomes based on sequence information, select and design primers for site validation and give visual representation of the event to guide downstream experiments.

**License** GPL-2

**LazyData** TRUE

**Depends** R (>= 3.3.0), grDevices, stats

**Imports** graphics, ggplot2, seqLogo, Biostrings, biomaRt, GenomicAlignments, GenomicRanges, GenomicFeatures, Gviz, IRanges, S4Vectors, GenomeInfoDb, utils, plyr, BSgenome.Mmusculus.UCSC.mm9

**biocViews** GeneExpression, RNASeq, Visualization, AlternativeSplicing

**Collate** primerpcr.R main\_splinter.R

**RoxygenNote** 6.0.1

**VignetteBuilder** knitr

**Suggests** BiocStyle, knitr, rmarkdown

**NeedsCompilation** no

## R topics documented:

acceptor.m . . . . .	2
addEnsemblAnnotation . . . . .	3
callPrimer3 . . . . .	4
checkPrimer . . . . .	5
compatible_cds . . . . .	5
compatible_tx . . . . .	6
donor.m . . . . .	6
eventOutcomeCompare . . . . .	7

eventOutcomeTranslate . . . . .	7
eventPlot . . . . .	8
extendROI . . . . .	9
extractSpliceEvents . . . . .	10
extractSpliceSites . . . . .	11
findCompatibleEvents . . . . .	12
findCompatibleExon . . . . .	13
findTX . . . . .	13
getPCRsizes . . . . .	14
getRegionDNA . . . . .	15
insertRegion . . . . .	15
makeROI . . . . .	16
makeUniqueIDs . . . . .	17
pcr_result1 . . . . .	17
plot_seqlogo . . . . .	18
primers . . . . .	18
psiPlot . . . . .	19
region_minus_exon . . . . .	20
removeRegion . . . . .	20
roi . . . . .	21
shapiroAcceptor . . . . .	21
shapiroDensity . . . . .	22
shapiroDonor . . . . .	23
splice_data . . . . .	24
splice_fasta . . . . .	24
splitPCRhit . . . . .	25
the cds . . . . .	25
the exons . . . . .	26
valid_cds . . . . .	26
valid_tx . . . . .	27
<b>Index</b>	<b>28</b>

---

acceptor.m

*acceptor.m*

---

## Description

Acceptor site mammalian frequency matrices for GT-AG pairs from SpliceDB

## Usage

```
data("acceptor.m")
```

## Format

The format is: num [1:4, 1:15] 9 31.03 12.5 42.36 8.44 ... - attr(\*, "dimnames")=List of 2 ..\$ : chr [1:4] "A" "C" "G" "T" ..\$ : chr [1:15] "V1" "V2" "V3" "V4" ...

## Source

[urlhttp://www.softberry.com/spldb/SpliceDB.html](http://www.softberry.com/spldb/SpliceDB.html)

## References

Burset M., Seledtsov I., Solovyev V. (Nucl. Acids Res., 2000, 28, 4364-4375; Nucl. Acids Res., 2001, 29, 255-259)

## Examples

```
data(acceptor.m)
```

---

addEnsemblAnnotation    *addEnsemblAnnotation*

---

## Description

Adds annotation to [extractSpliceEvents](#) object (if not present)

## Usage

```
addEnsemblAnnotation(data, species = "hsapiens")
```

## Arguments

data	<a href="#">extractSpliceEvents</a> object
species	character. biomaRt species passed to retrieve annotation. Common species include: 'hsapiens', 'mmusculus'

## Value

[extractSpliceEvents](#) object with annotated genes under \$geneSymbol

## Author(s)

Diana Low

## See Also

[http://asia.ensembl.org/info/data/biomart/biomart\\_r\\_package.html#biomartexamples](http://asia.ensembl.org/info/data/biomart/biomart_r_package.html#biomartexamples)

## Examples

```
data_path<-system.file("extdata",package="SPLINTER")
splice_data<-extractSpliceEvents(data=paste(data_path, "/skipped_exons.txt", sep=""))
splice_data<-addEnsemblAnnotation(data=splice_data,species="mmusculus")
```

---

 callPrimer3

*callPrimer3*


---

### Description

call primer3 for a given set of DNASTringSet object

### Usage

```
callPrimer3(seq, size_range = "150-500", Tm = c(57, 59, 62),
  name = "Primer1", primer3 = "primer3-2.3.7/bin/primer3_core",
  thermo.param = "primer3-2.3.7/src/primer3_config/",
  sequence_target = NULL,
  settings = "primer3-2.3.7/primer3web_v4_0_0_default_settings.txt")
```

### Arguments

seq	DNASTring object, one DNA string for the given amplicon
size_range	default: '151-500'
Tm	melting temprature parameters default:c(55,57,58)
name	name of the amplicon in chr_start_end format
primer3	primer3 path
thermo.param	thermodynamic parameters folder
sequence_target	If one or more targets is specified then a legal primer pair must flank at least one of them.
settings	text file for parameters

### Details

modified to include SEQUENCE\_TARGET as an option

### Value

data.frame of designed primers and parameters

### Author(s)

Altuna Akalin's modified Arnaud Krebs' original function further modified here by Diana Low

### Examples

```
### NOT RUN ###
# primer_results<-callPrimer3(seq='')
```

---

checkPrimer	<i>checkPrimer</i>
-------------	--------------------

---

**Description**

checkPrimer

**Usage**

```
checkPrimer(pp, genome, roi = NULL)
```

**Arguments**

pp	data.frame defining primers, or output of <a href="#">callPrimer3</a> . minimal columns = PRIMER_LEFT_SEQUENCE,PRIMER_RIGHT_SEQUENCE
genome	BSgenome object
roi	<a href="#">makeROI</a> object

**Value**

list of GRanges with primer locations

**Author(s)**

Diana Low

**Examples**

```
# create a primer pair
roi
primer_pair <- data.frame(PRIMER_LEFT_SEQUENCE="agctcttgaaattggagctgac",
                          PRIMER_RIGHT_SEQUENCE="cttagaaagaacaggaaatcc",
                          stringsAsFactors=FALSE)
```

---

compatible_cds	<i>compatible_cds</i>
----------------	-----------------------

---

**Description**

compatible\_cds

**Examples**

```
data(compatible_cds)
## maybe str(compatible_cds) ; plot(compatible_cds) ...
```

---

compatible_tx	<i>compatible_tx</i>
---------------	----------------------

---

### Description

compatible\_tx

### Examples

```
data(compatible_tx)
## maybe str(compatible_tx) ; plot(compatible_tx) ...
```

---

donor.m	<i>donor.m</i>
---------	----------------

---

### Description

Donor site mammalian frequency matrices for GT-AG pairs from SpliceDB

### Usage

```
data("donor.m")
```

### Format

The format is: num [1:4, 1:9] 34.1 36.2 18.3 11.4 60.4 ... - attr(\*, "dimnames")=List of 2 ..\$ : chr [1:4] "A" "C" "G" "T" ..\$ : chr [1:9] "V1" "V2" "V3" "V4" ...

### Source

<http://www.softberry.com/spldb/SpliceDB.html>

### References

Burset M., Seledtsov I., Solovyev V. (Nucl.Acids Res.,2000,28,4364-4375; Nucl. Acids Res.,2001,29,255-259)

### Examples

```
data(donor.m)
```

---

eventOutcomeCompare    *eventOutcomeCompare*

---

**Description**

Compares two sequences and gives differences if there's a switch from 1->2 if seq2 is NULL, assume seq1 is a list of length 2 to compare

**Usage**

```
eventOutcomeCompare(seq1, seq2 = NULL, genome, direction = TRUE,
  fullseq = TRUE)
```

**Arguments**

seq1	GRangesList
seq2	GRangesList
genome	BSGenome object
direction	logical. Report direction of sequence change.
fullseq	logical. Report full sequences.

**Value**

list containing  
 (1) tt : PairwiseAlignmentsSingleSubject pairwise alignment  
 (2) eventtypes : string detailing primary event classification

**Author(s)**

Diana LOW

**Examples**

```
suppressMessages(library(BSgenome.Mmusculus.UCSC.mm9))
bsgenome<-BSgenome.Mmusculus.UCSC.mm9
eventOutcomeCompare(seq1=compatible_cds$hits[[1]],seq2=region_minus_exon,
  genome=bsgenome,direction=TRUE)
```

---

eventOutcomeTranslate    *eventOutcomeTranslate*

---

**Description**

translates sequences, reports if NMD or NTC

**Usage**

```
eventOutcomeTranslate(seq1, genome, direction = FALSE, fullseq = TRUE)
```

**Arguments**

seq1	GRangesList
genome	BSGenome object
direction	logical. Report direction of sequence change.
fullseq	logical. Output full AA sequence.

**Value**

list of translated sequences

**Author(s)**

Diana LOW

**Examples**

```
suppressMessages(library(BSgenome.Mmusculus.UCSC.mm9))
bsgenome<-BSgenome.Mmusculus.UCSC.mm9
translation_results<-eventOutcomeTranslate(compatible_cds,genome=bsgenome,
direction=TRUE)
```

---

eventPlot

*eventPlot*

---

**Description**

eventPlot

**Usage**

```
eventPlot(transcripts, roi_plot = NULL, bams = c(), names = c(),
annoLabel = c("Gene A"), rspan = 1000, showAll = TRUE)
```

**Arguments**

transcripts	GRanges object
roi_plot	GRanges object region to plot
bams	character vector of bam file locations
names	character vector of name labels
annoLabel	character. annotation label
rspan	integer or NULL. number of basepairs to span from roi. if NULL, will consider whole gene of roi
showAll	logical. TRUE = display splice junctions of entire view or FALSE = just roi.

**Value**

a Gviz plot of genomic region



**Author(s)**

Diana Low

**Examples**

```
# define BAM files
data_path<-system.file("extdata",package="SPLINTER")
mt<-paste(data_path,"/mt_chr14.bam",sep="")
wt<-paste(data_path,"/wt_chr14.bam",sep="")

# plot results
eventPlot(transcripts=valid_tx,roi_plot=roi,bams=c(wt,mt),
  names=c('wt','mt'),rspan=1000)
```

---

`extendROI`*extendROI*

---

**Description**

extend the span of the current ROI by n number of up/downstream exon(s) by modifying `roi_range` within the `makeROI` object while retaining legacy sites by keeping `$roi` and `$flank`

**Usage**

```
extendROI(roi, tx, up = 0, down = 0, type = 1)
```

**Arguments**

<code>roi</code>	<code>makeROI</code> object
<code>tx</code>	<code>GRangesList</code> transcript list to pull regions from
<code>up</code>	integer. number of exons to extend upstream
<code>down</code>	integer. number of exons to extend downstream
<code>type</code>	integer. 1=full cassette, 2=flank only

**Value**

`makeROI` object with modified ranges

**Examples**

```
extendROI(roi,valid_tx,up=1)
```

---

extractSpliceEvents     *extractSpliceEvents*

---

### Description

Extracts the location of target, upstream and downstream splice sites Used for calculations and genome visualizations

### Usage

```
extractSpliceEvents(data = NULL, filetype = "mats", splicetype = "SE",
  fdr = 1, inclusion = 1, start0 = TRUE)
```

### Arguments

data	character. path to file
filetype	character. type of splicing output. c('mats','custom'). see Details.
splicetype	character. c('SE', 'RI', 'MXE', 'A5SS', 'A3SS')
fdr	numeric. false discovery rate filter range [0,1]
inclusion	numeric. splicing inclusion range, takes absolute value
start0	boolean 0-base start

### Details

filetype 'custom' should provide a 9-column tab-delimited text file with the following columns: ID (Ensembl gene id), Symbol (gene name), chr, strand, exonStart, exonEnd, exon2Start, exon2End, upstreamStart, upstreamEnd, downstreamStart, downstreamEnd eg. ENSG0000012345 chr1 + 3 4 5 6 1 2 7 8

### Value

list containing information on

- (1) original file type
- (2) splice event type
- (3) data.frame with splicing regions

### Author(s)

Diana Low

### See Also

[http://rnaseq-mats.sourceforge.net/user\\_guide.htm](http://rnaseq-mats.sourceforge.net/user_guide.htm) for MATS file definition

### Examples

```
data_path<-system.file("extdata",package="SPLINTER")
splice_data<-extractSpliceEvents(data=paste(data_path,"/skipped_exons.txt",sep=""))
```

---

extractSpliceSites      *extractSpliceSites*

---

## Description

Extracts and formats to bed the location of target, upstream and downstream splice sites

## Usage

```
extractSpliceSites(df, target = "SE", site = "donor", motif_range = c(-3,
6), start0 = TRUE)
```

## Arguments

df	extractSpliceEvents object
target	the target site to extract. See Details.
site	character donor or acceptor
motif_range	numeric vector of splice position to extract
start0	boolean 0-base start

## Details

target : the site to extract the sequence from. It can be either the event in question (SE, RI, MXE - first exon, MXE2 - second exon, A5SSlong, A5SSshort, A3SSlong, A3SSshort, upstream or downstream). If this function is used in conjunction with [shapiroDonor](#) or [shapiroAcceptor](#) to compute scores, then most likely it will be run twice - once for the event, and the other either up- or downstream as a comparison.

## Value

GRanges object

## Author(s)

Diana Low

## See Also

[http://rnaseq-mats.sourceforge.net/user\\_guide.htm](http://rnaseq-mats.sourceforge.net/user_guide.htm) for MATS file definition

## Examples

```
data_path<-system.file("extdata",package="SPLINTER")
splice_data<-extractSpliceEvents(data=paste(data_path,"/skipped_exons.txt",sep=""))
splice_sites<-extractSpliceSites(splice_data,target="SE")
```

---

findCompatibleEvents *findCompatibleEvents*

---

### Description

Which transcript contains the event? Each event has 2 possibilities, as long as the transcript fulfills one, it passes the test Has to be exact (inner junctions)

### Usage

```
findCompatibleEvents(tx, tx2 = NULL, roi, sequential = TRUE,  
  verbose = FALSE)
```

### Arguments

tx	GRangesList object of transcripts
tx2	optional GRangesList object of transcripts if tx is list of cds
roi	<a href="#">makeROI</a> object containing event information
sequential	logical. Exons have to appear sequentially to be considered compatible
verbose	logical. printouts and messages.

### Details

Separates into event/region1 and 2 for the alternative case

### Value

list of length 4  
(1) GRangesList  
(2) Hits status [c]=coding; [nc]=non-coding  
(3) ct - compatible transcripts  
(4) tt - total transcripts

### Author(s)

Diana Low

### Examples

```
compatible_cds <- findCompatibleEvents(valid_cds,roi=roi,verbose=TRUE)
```

---

```
findCompatibleExon    findCompatibleExon
```

---

**Description**

Finds compatible exon in annotation with the one present in roi object

**Usage**

```
findCompatibleExon(tx, roi, verbose = FALSE)
```

**Arguments**

tx	GRangesList object of transcripts
roi	<a href="#">makeROI</a> object containing event information
verbose	logical. printouts and messages.

**Value**

list of length 3  
 (1) GRangesList hits  
 (2) Number of transcripts  
 (3) Original number of input transcripts

**Author(s)**

Diana Low

**Examples**

```
compatible_exons <- findCompatibleExon(valid_cds,roi)
```

---

```
findTX                findTX
```

---

**Description**

Given an ENSEMBL id, find all transcripts that matches id

**Usage**

```
findTX(id, db, tx, valid = FALSE)
```

**Arguments**

id	character. transcript identification (currently ENSEMBL gene names)
db	TxDb object
tx	GRangesList
valid	logical. check if in multiples of 3 [TRUE] for CDS translation.

**Value**

GRangesList

**Author(s)**

Diana Low

**Examples**

```
valid_cds <- findTX(id=splice_data$data[1,]$GeneID, tx=the cds, db=txdb, valid=FALSE)
```

---

getPCRsizes

*getPCRsizes*


---

**Description**

returns length of product given a GRanges span and GRangesList of transcripts

**Usage**

```
getPCRsizes(pcr_span, txlist, verbose = FALSE)
```

**Arguments**

pcr_span	GRanges object
txlist	GRangesList object
verbose	logical. report intermediate output.

**Value**

data.frame of transcript names with detected sizes in basepairs

**Author(s)**

Diana Low

**Examples**

```
suppressMessages(library(BSgenome.Mmusculus.UCSC.mm9))
bsgenome<-BSgenome.Mmusculus.UCSC.mm9
## create a primer pair
## for actual use, obtain primer pair from primer design (callPrimer3)
primer_pair <- data.frame(PRIMER_LEFT_SEQUENCE="agctcttgaaattggagctgac",
                          PRIMER_RIGHT_SEQUENCE="cttagaagaacaggaaatcc",
                          stringsAsFactors=FALSE)

## confirm location
cp<-checkPrimer(primer_pair,bsgenome,roi)
cp

## get the PCR sizes
pcr_result1 <- getPCRsizes(cp,theexons)
```

---

getRegionDNA	<i>getRegionDNA</i>
--------------	---------------------

---

**Description**

get DNA sequence give a region of interest

**Usage**

```
getRegionDNA(roi, genome, introns = FALSE)
```

**Arguments**

roi	<a href="#">makeROI</a> object
genome	BSgenome object
introns	TRUE/FALSE. whether to include intronic (lowercase) DNA. By default returns only exonic (uppercase) DNA.

**Value**

list of  
 (1) DNA sequence (2) Junction start (for primer design)

**Author(s)**

Diana Low

**Examples**

```
suppressMessages(library(BSgenome.Mmusculus.UCSC.mm9))
bsgenome<-BSgenome.Mmusculus.UCSC.mm9
getRegionDNA(roi,bsgenome)
```

---

insertRegion	<i>insertRegion</i>
--------------	---------------------

---

**Description**

inserts a region (exon or intron) into roi object

**Usage**

```
insertRegion(subject, roi)
```

**Arguments**

subject	GrangesList
roi	<a href="#">makeROI</a> object containg region of interest (to insert). refer to makeROI().

**Details**

in the case of intron retention, replaces exon with intron retention range reduce() the GRanges in question

**Value**

GRanges object

**Author(s)**

Diana Low

**Examples**

```
#Inserts the exon defined in roi GRanges object from a GRanges/GRangesList
region_minus_exon
region_with_exon<-insertRegion(region_minus_exon,roi)
```

---

makeROI

*makeROI*

---

**Description**

Creates an object to store information about the splice site (region of interest) including flanking regions and alternative splice outcome

**Usage**

```
makeROI(df, type = "SE")
```

**Arguments**

df                    data.frame object from [extractSpliceEvents](#)  
 type                  type of splicing event c("SE","RI","MXE","A5SS","A3SS")

**Value**

a list containing

- (1) type : splice type
- (2) name : ID of transcript
- (3) roi : GRanges object of splice site
- (4) flank : GRanges object of flanking exons of splice site
- (5) roi\_range : GRangesList of splice site and it's alternative outcome based on type

**Author(s)**

Diana Low

**Examples**

```
single_record<-splice_data$data[which(grepl("Prmt5",splice_data$data$Symbol)),]
roi <- makeROI(single_record,type="SE")
```



---

<code>makeUniqueIDs</code>	<i>makeUniqueIDs</i>
----------------------------	----------------------

---

**Description**

Makes unique ID names from event location

**Usage**

```
makeUniqueIDs(data)
```

**Arguments**

`data`                    `extractSpliceEvents` object

**Value**

original `extractSpliceEvents` list object with unique ID appended to data accessor

**Author(s)**

Diana Low

**Examples**

```
data_with_id<-makeUniqueIDs(splice_data)
```

---

<code>pcr_result1</code>	<i>pcr_result1</i>
--------------------------	--------------------

---

**Description**

`pcr_result1`

**Examples**

```
data(pcr_result1)
```

plot\_seqlogo                    *plotting sequence logo*

---

**Description**

Plots the sequence logo of a given set of FASTA sequences

**Usage**

```
plot_seqlogo(fasta_seq)
```

**Arguments**

fasta\_seq            DNASTringSet or path to fasta-formatted file

**Value**

sequence logo image

**Author(s)**

Diana Low

**Examples**

```
head(splice_fasta)
test<-Biostrings::DNASTringSet(splice_fasta$V2)
plot_seqlogo(test)
```

---

primers                    *primers*

---

**Description**

primers designed using Primer3 for sample data

**Usage**

```
data("primers")
```

**Format**

A data frame with 5 observations on the following 28 variables.

i a numeric vector

PRIMER\_LEFT\_SEQUENCE a character vector

PRIMER\_RIGHT\_SEQUENCE a character vector

PRIMER\_LEFT\_TM a numeric vector

PRIMER\_RIGHT\_TM a numeric vector

PRIMER\_LEFT\_pos a numeric vector  
 PRIMER\_LEFT\_len a numeric vector  
 PRIMER\_RIGHT\_pos a numeric vector  
 PRIMER\_RIGHT\_len a numeric vector  
 PRIMER\_PAIR\_PENALTY a numeric vector  
 PRIMER\_LEFT\_PENALTY a numeric vector  
 PRIMER\_RIGHT\_PENALTY a numeric vector  
 PRIMER\_LEFT\_GC\_PERCENT a numeric vector  
 PRIMER\_RIGHT\_GC\_PERCENT a numeric vector  
 PRIMER\_LEFT\_SELF\_ANY\_TH a numeric vector  
 PRIMER\_RIGHT\_SELF\_ANY\_TH a numeric vector  
 PRIMER\_LEFT\_SELF\_END\_TH a numeric vector  
 PRIMER\_RIGHT\_SELF\_END\_TH a numeric vector  
 PRIMER\_LEFT\_HAIRPIN\_TH a numeric vector  
 PRIMER\_RIGHT\_HAIRPIN\_TH a numeric vector  
 PRIMER\_LEFT\_END\_STABILITY a numeric vector  
 PRIMER\_RIGHT\_END\_STABILITY a numeric vector  
 PRIMER\_LEFT\_TEMPLATE\_MISPRIMING a numeric vector  
 PRIMER\_RIGHT\_TEMPLATE\_MISPRIMING a numeric vector  
 PRIMER\_PAIR\_COMPL\_ANY\_TH a numeric vector  
 PRIMER\_PAIR\_COMPL\_END\_TH a numeric vector  
 PRIMER\_PAIR\_PRODUCT\_SIZE a numeric vector  
 PRIMER\_PAIR\_TEMPLATE\_MISPRIMING a numeric vector

**Value**

Dataframe of primer design results

**Examples**

```
data(primers)
```

---

 psiPlot

*psiPlot*


---

**Description**

Plots percentage spliced in (PSI) values in terms of inclusion levels

**Usage**

```
psiPlot(df = NULL, type = "MATS", sample_labels = c("Sample 1",
  "Sample 2"))
```

**Arguments**

df                    data.frame containing PSI values  
 type                character. either 'MATS' output (will read in MATS headers) or 'generic' (provide 4 or 6 column data.frame)  
 sample\_labels      x-axis labels for the plot

**Value**

bar plot of PSI values

**Author(s)**

Diana Low

**Examples**

```
#we give inclusion and skipped numbers as reads
#this will be converted into percentages
df<-data.frame(inclusion1=c("6,4,6"),skipped1=c("10,12,12"),inclusion2=c("15,15,15"),
               skipped2=c("3,3,4"),stringsAsFactors = FALSE)
psiPlot(df,type='generic')
```

---

region_minus_exon	<i>region_minus_exon</i>
-------------------	--------------------------

---

**Description**

region\_minus\_exon

**Examples**

```
data(region_minus_exon)
## maybe str(region_minus_exon) ; plot(region_minus_exon) ...
```

---

removeRegion	<i>removeRegion</i>
--------------	---------------------

---

**Description**

removes a region (exon) from a GRanges or GRangesList

**Usage**

```
removeRegion(subject, roi)
```

**Arguments**

subject            GRanges or GrangesList object  
 roi                [makeROI](#) object containing GRanges range (to remove)

**Value**

GRanges object

**Author(s)**

Diana Low

```
# Removes the exon defined in roi GRanges object from a GRanges/GRangesList compatible_cds$hits[[1]]
region_minus_exon<-removeRegion(compatible_cds$hits[[1]],roi)
```

---

roi	<i>roi</i>
-----	------------

---

**Description**

roi

**Usage**

```
data("roi")
```

**Value**

List containing region of interest information

**Examples**

```
data(roi)
```

---

shapiroAcceptor	<i>shapiroAcceptor</i>
-----------------	------------------------

---

**Description**

Shapiro's score of acceptor site (range is from -13 [intron] to +1 [exon]) is:  $100 * ((t1 - l1)/(h1 - l1) + (t2 - l2)/(h2 - l2))/2$ , where t1 is the sum of the best 8 of 10 percentages at positions -13 to -4, l1 is the sum of the lowest 8 of 10 percentages at position -13 to -4, h1 is the sum of the highest 8 of 10 percentages at positions -13 to -4, t2 is the sum of percentages at positions -3 to +1, l2 is the sum of the lowest percentages at positions -3 to +1, and h2 is the sum of the highest percentages at positions -3 to +1

**Usage**

```
shapiroAcceptor(reference_fasta, target_fasta)
```

**Arguments**

reference_fasta	vector of strings or DNASTringSet of reference splice list
target_fasta	vector of strings or DNASTringSet of fasta to score

**Value**

data.frame with Shapiro scores

**Author(s)**

Diana Low

**See Also**

<http://www.softberry.com/spldb/SpliceDB.html>

**Examples**

```
library(BSgenome.Mmusculus.UCSC.mm9)
bgenome <- BSgenome.Mmusculus.UCSC.mm9
data_path<-system.file("extdata",package="SPLINTER")
splice_data<-extractSpliceEvents(data=paste(data_path,"/skipped_exons.txt",sep=""))
splice_sites<-extractSpliceSites(splice_data,site="acceptor")
acceptor.ss<-getSeq(bgenome,splice_sites)
sacceptor<-shapiroAcceptor(acceptor.m,acceptor.ss)
```

---

shapiroDensity

*shapiroDensity*

---

**Description**

convenience function for plotting Shapiro score density

**Usage**

```
shapiroDensity(ctrl_scores, treat_scores, sample = c(1, 2))
```

**Arguments**

ctrl_scores	output of shapiroDonor or shapiroAcceptor
treat_scores	output of shapiroDonor or shapiroAcceptor
sample	samplenames

**Author(s)**

Diana Low

---

shapiroDonor	<i>shapiroDonor</i>
--------------	---------------------

---

## Description

Shapiro and Senapathy (1987) have developed a method to score the strength of a splice site based on percentages of each nucleotide at each position. Shapiro's score of donor site (range is from -3 [exon] to +7 [intron]) is :  $100 * (t - \text{min}) / (\text{max} - \text{min})$ , where t is the sum of percentages at positions -3 to +7, min is the sum of the lowest percentages at positions -3 to +7, and max is the sum of the highest percentages at positions -3 to +7.

## Usage

```
shapiroDonor(reference_fasta, target_fasta)
```

## Arguments

```
reference_fasta      vector of strings or DNASTringSet of reference splice list
target_fasta        vector of strings or DNASTringSet of fasta to score
```

## Value

data.frame with Shapiro scores

## Author(s)

Diana Low

Diana Low

## See Also

<http://www.softberry.com/spldb/SpliceDB.html>

## Examples

```
library(BSgenome.Mmusculus.UCSC.mm9)
bsgenome <- BSgenome.Mmusculus.UCSC.mm9
data_path<-system.file("extdata",package="SPLINTER")
splice_data<-extractSpliceEvents(data=paste(data_path,"/skipped_exons.txt",sep=""))
splice_sites<-extractSpliceSites(splice_data)
donor.ss<-getSeq(bsgenome,splice_sites)
sdonor<-shapiroDonor(donor.m,donor.ss)
```

---

splice_data	<i>splice_data</i>
-------------	--------------------

---

**Description**

splice\_data

**Usage**

```
data("splice_data")
```

**Value**

List containing splice event file information

**Examples**

```
data(splice_data)
```

---

splice_fasta	<i>splice_fasta</i>
--------------	---------------------

---

**Description**

splice\_fasta

**Usage**

```
data("splice_fasta")
```

**Format**

A data frame with 0 observations on the following 2 variables.

V1 a numeric vector

V2 a numeric vector

**Value**

Dataframe of region and fasta sequence

**Examples**

```
data(splice_fasta)
```



---

splitPCRhit	<i>splitPCRhit</i>
-------------	--------------------

---

**Description**

splits the PCR alignment into the two AS conditions

**Usage**

```
splitPCRhit(res, hitlist)
```

**Arguments**

res	result from <a href="#">getPCRsizes</a>
hitlist	<a href="#">findCompatibleEvents</a> object

**Value**

list of 2 data.frame objects with isoform name (ID) and length of PCR product (bp) matching Type 1 or Type 2 transcripts

**Author(s)**

Diana Low

**Examples**

```
## as getPCRsizes gives you all PCR bands when the primers are used,  
## splitPCRhit will determine which bands are relevant to the target  
relevant_pcr_bands<-splitPCRhit(pcr_result1,compatible_tx)
```

---

the cds	<i>the cds</i>
---------	----------------

---

**Description**

the cds

**Usage**

```
data("the cds")
```

**Value**

List containing GRanges info

**Examples**

```
data(the cds)
```

---

theexons	<i>theexons</i>
----------	-----------------

---

**Description**

theexons

**Usage**

```
data("theexons")
```

**Value**

List containing GRanges info

**Examples**

```
data(theexons)
```

---

valid_cds	<i>valid_cds</i>
-----------	------------------

---

**Description**

valid\_cds

**Usage**

```
data("valid_cds")
```

**Value**

GRangesList

**Examples**

```
data(valid_cds)
```

---

*valid\_tx*

*valid\_tx*

---

**Description**

*valid\_tx*

**Value**

GRangesList

**Examples**

```
data(valid_tx)
## maybe str(valid_tx) ; plot(valid_tx) ...
```

# Index

## \*Topic **datasets**

acceptor.m, 2  
compatible\_cds, 5  
compatible\_tx, 6  
donor.m, 6  
pcr\_result1, 17  
primers, 18  
region\_minus\_exon, 20  
roi, 21  
splice\_data, 24  
splice\_fasta, 24  
thecds, 25  
theexons, 26  
valid\_cds, 26  
valid\_tx, 27

acceptor.m, 2  
addEnsemblAnnotation, 3

callPrimer3, 4, 5  
checkPrimer, 5  
compatible\_cds, 5  
compatible\_tx, 6

donor.m, 6

eventOutcomeCompare, 7  
eventOutcomeTranslate, 7  
eventPlot, 8  
extendROI, 9  
extractSpliceEvents, 3, 10, 16  
extractSpliceSites, 11

findCompatibleEvents, 12, 25  
findCompatibleExon, 13  
findTX, 13

getPCRsizes, 14, 25  
getRegionDNA, 15

insertRegion, 15

makeROI, 5, 9, 12, 13, 15, 16, 20  
makeUniqueIDs, 17

pcr\_result1, 17

plot\_seqlogo, 18  
primers, 18  
psiPlot, 19

region\_minus\_exon, 20  
removeRegion, 20  
roi, 21

shapiroAcceptor, 11, 21  
shapiroDensity, 22  
shapiroDonor, 11, 23  
splice\_data, 24  
splice\_fasta, 24  
splitPCRhit, 25

thecds, 25  
theexons, 26

valid\_cds, 26  
valid\_tx, 27